



dbVar: A Genomic Structural Variation Database

A collection of human genomic structural variations

<https://www.ncbi.nlm.nih.gov/dbvar/>

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

Scope and Access

The NCBI dbVar Structural Variation database houses human genomic structural variants (SV) greater than 50 base pairs in length. From the dbVar homepage (A, www.ncbi.nlm.nih.gov/dbvar/), you can search (B), browse (C), view and download variant data from over 150 studies, such as 1000 Genomes Phase 3 (estd219), Simons Genome Diversity Project (nstd128), ClinGen (nstd45), and ExAC (nstd151). You can access the variants using the Study Browser or the graphical Genome Browser. Individual Study and Variant Pages include links to raw data as well as to related information at other NCBI and external resources. Bulk data downloads are available by FTP at <ftp.ncbi.nlm.nih.gov/pub/dbVar/data> (D).



In 2018, dbVar introduced a new comprehensive set of non-redundant structural variants (NR set) consisting of unique insertions, duplications, and deletions. These compact files are suitable for use as references in the analysis of human structural variation. For example, you can use the dbVar NR set to filter and annotate other datasets in a broad range of applications including variant discovery and identifying rare and/or clinical variants. The dbVar NR set currently includes more than 2.2 million deletions, 1.1 million insertions, and 300 thousand duplications. The NR set will be updated monthly as new variants are added to dbVar. You can find more information about NR set, brief tutorials, and ways to access NR set FTP files from github at github.com/ncbi/dbvar/tree/master/Structural_Variant_Sets (E).

Searching in dbVar

To search dbVar, type terms in the search box and click the "Search" button (F). For more refined results, use field-limited terms connected with Boolean operators. For example, querying with [human\[orgn\]](#) [AND 17\[chromosome\]](#) [AND deletion\[phenotype\]](#), retrieves structural deletion variants on human chromosome 17 (G).

The default display lists the variant ID, the type of variation and other key features in a table. Click the coordinates on different genome assemblies in the Location column (H) to open a graphical presentation of the variants in the dbVar Genome Browser. Click the "... more genes" link in the "Genes in region" column (I) to change display to the Summary format so affected gene records are directly accessible. Use the Study filter (J) in the left-hand column to access relevant studies.

The screenshot shows the dbVar homepage with various navigation links and a search bar. The search results page displays a table of variants with columns for Variant Region ID, Type, Number of Variant Calls, Study ID, Organism, Clinical Assertion, Location, and Genes in region. The variant details page shows information about a specific variant, including its coordinates, study, and clinical significance.

dbVar
dbVar is NCBI's database of human genomic structural variation — insertions, deletions, duplications, inversions, mobile elements, and translocations

Getting Started
[Overview of Structural Variation](#)
[Organism List](#)
[FAQ](#)
[Help](#)
[Factsheet](#)

Accessing Data
[Structural Variation Data Hub](#)
[Tools for analyzing dbVar data](#)
[Study Browser](#) (C)
[Genome Browser](#) (D)
[FTP Data Download](#)

Other NCBI Resources
[dbSNP](#)
[ClinVar](#)
[Variation Portal](#)
[Variation Tools](#)

Non-Redundant Structural Variation Datasets
Would you like to compare and analyze your data with known structural variants (SV) in dbVar? Now there are easy-to-use files containing non-redundant (NR) deletions, duplications, and insertions aggregated from across studies in dbVar. The files are available for human assembly versions GRCh37 and GRCh38. Descriptions of the NR data and tutorials are available on [GitHub](#).

Submitting Data
[Submission Guidelines](#)
[Submission Templates](#)
[VCF Submissions](#)

dbVar News
[Announcements](#)

External Resources
[Database of Genomic Variants archive \(DGVa\)](#)
[1000 Genomes Project](#)

Search results
Items: 8

Number of Variants: 8

Variant Region ID	Type	Number of Variant Calls	Study ID	Organism	Clinical Assertion	Location	Genes in region
nsv1398520	copy number variation	1	nstd102	human	Pathogenic	NCBI36 (hg18) chr17: 6,936,148-7,177,360 , GRCh37 (hg19) chr17: 6,995,424-7,236,636 , GRCh38 (hg38) chr17: 7,092,105-7,333,317	BARAP, CLDN7, 14 more genes
nsv1398446	copy number variation	1	nstd102	human	Pathogenic	NCBI36 (hg18) chr17: 6,996,378-7,152,828 , GRCh37 (hg19) chr17: 7,055,654-7,212,104 , GRCh38 (hg38) chr17: 7,152,335-7,308,785	GABARAP, CLDN7, 11 more genes

Summary
20 per page
Search results
Items: 8

1. Variant type: copy number variation
Associated study: [nstd102](#)
Organism: human
Genes(s) in region: [ACADVL](#), [ASGR1](#), [ASGF1](#)
[See more...](#)
Location information:
Submitted: NCBI36 (hg18): 17:6,936,148-7,177,360
Remapped: GRCh37 (hg19): 17:6,995,424-7,236,636
GRCh38 (hg38): 17:7,092,105-7,333,317
Validation status: Not tested
Clinical significance: Pathogenic
ID: 30348183 variant

Object Type
Variant (8)

Organism
human (8)

Variant Region Type
copy number variation (8)

Method Type
Multiple (5)
Curated (2)
Sequencing (1)

Clinical Assertion
Pathogenic (7)
Likely pathogenic (1)

Subject Gender
not reported (8)

Study Type
Case-Set (6)

Tabular View 20 per page
<http://bit.ly/2cSSByR>

Send to:

The Advanced Page

On the home page, clicking on the “Advanced” link beneath the search box takes you to the Advanced search page (A), which provides access to indexing fields (B) and terms indexed in each field (C) through the “Show index list” link (D), as well as your search history (E). The query builder function allows you to combine indexing terms and entries in the current search history for more specific retrieval. Click the “Edit” link (F) to unlock the query builder box so you can manually edit search terms.

- Chromosome Outer Start
- Chromosome Start
- ClinVar Accession
- Detection Method
- Filter**
- Gender
- Gene Full Name
- Gene Name
- Library Abbreviation
- MIM ID
- MeSH Terms
- MeSH Unique ID
- Method Platform

dbVar Advanced Search Builder (A)

Search box: "diabetes"[Phenotype] AND #21

Buttons: Cancel, Search, or Add to history, Clear

Builder section:

Filter: "clin pathogenic"[Filter] (C)

Index list (D):

- clin none (4470200)
- clin not provided (2296)
- clin pathogenic (5638)**
- clin risk factor (2)
- clin uncertain significance (4795)
- dbvar all (4481898)
- dbvar bioproject (617317)

Buttons: Hide index list, Previous 200, Next 200, Refresh index, Show index list

Search: Search or Add to history

History (E):

Search	Add to builder	Query	Items found	Time
#23	Add	Search "diabetes"[Phenotype] AND #21	9	16:16:49
#22	Add	Search "diabetes"[Phenotype]	192	16:15:27
#21	Add	Search "clin pathogenic"[Filter]	5638	16:14:02
#17	Add	Search human[orgn] AND 17[chromosome] AND deletion[phenotype]	4	16:02:16

dbVar Study Browser

The dbVar homepage provides links to the Study Browser page, where you can browse by study IDs. The Study Browser table (below) sums up the available studies in a table. Clicking a column header, such as the “Variant Region Count” (G), sorts the list according to the value in that column so you can locate studies with certain characteristics. Identifiers for individual studies in the Study column (H) links to detailed display of specific studies, while the citations in the Publication column (I) link to relevant articles in PubMed. You can use criteria filters in the right-hand column (J) to narrow the list of studies displayed.

dbVar: Study Browser

Date	Publication	Study	Organism	Variant Region Count	Variant Call Count
2018/09	Szafranski et al. 2018	nstd153	Human	8	8
2018/04	Kucukkilic et al. 2018	nstd159	Human	2	350
2018/02	Fu et al. 2018	nstd156	Human	963	11,796
2018/02	Adewoye et al. 2018	nstd155	Human	1	2,502
2018/01	Rambo-Martin et al. 2018	nstd141	Human	123	274
2017/12	Möller et al. 2017	nstd154	Human	1	7
2017/09	Gambin et al. 2017	nstd149	Human	36	41
2017/08	Gardner et al. 2017	nstd144	Chimpanzee, Human	37,798	38,422
2017/07	Luo et al. 2017b	estd233	Human	1,026	1,026
2017/07	Lu et al. 2017	nstd145	Human	8,237	26,917
2017/03	dbSNP integrated variants	nstd90	Human	4,715	4,722
2017/02	Blanco-Rodriguez et al. 2017	estd232	Human	8	8
2017/02	Shang et al. 2017	nstd143	Human	18	18
2017/02	Rahbari et al. 2017	nstd142	Human	3	7
2017/02	Walker et al. 2017	nstd132	Human	6,173	15,888
2017/01	Fan et al. 2017	nstd140	Human	22,644	22,644

Filter by Study Type

- Control Set (87)
- Case-Set (35)
- Case-Control (17)
- Collection (11)
- Curated Collection (5)
- See more... Source: NCBI

Filter by Method

- Sequencing (43)
- SNP array (31)
- Oligo aCGH (17)
- BAC aCGH (6)
- Curated (5)
- See more... Source: NCBI

Filter by Variant

- >=10

dbVar Genomes Browser

Ideogram View

Unplaced/unlocalized scaffolds: 159
Alt loci/patches: 349

1 2 3 4 5 6 7 8 9 10 11 12 13 14
15 16 17 18 19 20 21 22 X Y MT

Pick Assembly

GCF_000001405.34 (GRCh38.p8)

Select an assembly to change

Search

Q: TRT

Enter a location, gene name or phrase

Genes Other features

Gene	Location
TERT	Chr5: 1,253,167 - 1,295,047
AR	ChrX: 67,544,032 - 67,730,618
TRT-AGT5-1	Chr17: 8,139,452 - 8,139,525
TRT-AGT4-1	Chr8: 27,726,894 - 27,726,767
TRT-TGT6-1	Chr5: 181,191.7K - 181,191.8K
TRT-AGT2-2	Chr8: 27,684,695 - 27,684,768
TRT-TGT3-1	Chr14: 20,613,790 - 20,613,861

User Data and Track Hubs

Region Content

Data in view Click (-) to remove track

Study ID	Variant Calls
nstd54	32 (-)

Data available for region Click (+) to add track

Study ID	Variant Calls
estd3	1 (+)
estd20	7 (+)
estd22	19 (+)
estd59	118 (+)

dbVar Genome Browser

The Genome Browser (A) provides a way to locate reported variants for genomic regions of interests. You can specify the genomic regions by coordinates or annotated features known to occur in these regions, such as TRT (B). Click a retrieved entry from the list to zoom to the feature (C), click the (+) sign (D) to the right of studies to add tracks for mapped variants. Hover a variant to see its detail in a popup (E).

Homo sapiens: GRCh38.p8 (GCF_000001405.34) Chr 5 (NC_000005.10): 1,248,887 - 1,298,112

Reset All Share this page Help Version 2.7.5 Setup Page

NC_000005.10

Genes

Genomic Alignments

Repeat region

dbVar Cooper et al. 2011 (nstd54) Case Samples

ns537564 (+)

ns537572 (+)

ns537576 (+)

ns537579 (+)

ns537571 (+)

ns537578 (+)

ns537599 (+)

ns537605 (+)

ns537575

Variation ID: dbVar: ns537575

Supporting Variant Calls: 1

Variant Region Type: Copy number variation

Validation Status: Not Tested

Total Length: 27,442,368

Links & Tools

BLAST Genome-specific: NC_000005.10 (79,030..27,521,397)

BLAST Genomic: NC_000005.10 (79,030..27,521,397)

FASTA View: NC_000005.10 (79,030..27,521,397)

GenBank View: NC_000005.10 (79,030..27,521,397)

dbVar: ns537575

dbVar Study Pages

The details of a specific study are shown in the Study Page (F). The default "Variant Summary" tab (G) lists variants mapped to different chromosomes. Variants identified by the study are available through the "Variant in this study" link (H).

The "Samplesets" tab (I) displays samples used in the study. Information provided include the ID of the participating subject and their demonstrated phenotypes (J). The "Experimental Details" tab (K) provides methodologies and experimental platforms used when such information is made available by the submitters.

nstd54 (Cooper et al. 2011)

Organism: Human

Study Type: Case-Control

Submitter: Evan Eichler

Description: Copy Number Variants from 15,767 cases of Developmental Delay and Intellectual Disability from Signature Genomics, and 8329 Control Samples. This study contains samples in common with Coe et al. 2014. Due to analysis differences (see manuscripts) please use the case samples (Sampleset 1) from only one of these submissions. Control sample sets do not overlap and may be combined.

Publication(s): Cooper et al. 2011

Detailed Information: [Download 81345 Variant Regions](#), [Download 468909 Variant Calls](#), [Download Both](#), [FTP](#)

Variant Summary **G** Implesets Experimental Details Validations

Assembly used for analysis:

Remapped: GRCh38.p2 (hg38)

Remapped: GRCh37.p13 (hg19)

Submitted: NCBI36 (hg18)

Variant Summary for: GRCh38.p2 (hg38)

Sequence ID	Chr	Number of Variant Regions	Number of Variant Calls	Placement type	Link to graphical display
NC_000001.11	Chr1	5,611	34,158	Remapped	NC_000001.11
NC_000002.12	Chr2	5,509	35,198	Remapped	NC_000002.12
NC_000003.12	Chr3	4,584	31,407	Remapped	NC_000003.12

Variant Summary **I** Samplesets **K** Experimental Details Validations

Number of Samplesets: 12

Sampleset ID: 1

Sampleset Type: Case

Description: Children referred to Signature Genomics with a general diagnosis of intellectual disability (ID) and/or developmental delay (DD).

Size: 15,767

Organisms: [Homo sapiens](#)

Sampleset Phenotype(s): [Developmental Disabilities](#)

[Download Samples as CSV file](#)

Samples for sampleset 1 (displaying 100 of the 15767 samples)

Sample ID	Subject ID	Subject Phenotype
9881096	9881096	Developmental Disabilities Global developmental delay
9881101	9881101	Autism Developmental Disabilities Global developmental delay
9881113	9881113	Abnormality of the nervous system Developmental Disabilities Global developmental delay
9881122	9881122	Abnormality of the nervous system Developmental Disabilities Global developmental delay

dbVar Variant Record View

In the tabular display of a search result (A), clicking the variant ID (B) opens a detailed variant view display of that record.

The top section of this variant view (C), from left to right, provides a summary of the variant, overview of its genomic placement, and links to relevant records in other NCBI databases.

nsv997234 (C)

Organism: [Homo sapiens](#)
 Study: [nstd45 \(ClinGen Curated Dosage Sensitivity Map\)](#) Variant Calls: 1
 Variant Type: copy number variation Validation: Not tested
 Method Type: Curated Clinical Assertions: [Yes](#)
 Submitted on: GRCh37 (hg19) Region Size: 282,751
 Description: NF1
 Publication(s): [Estivill et al. 1991](#), [Riggs et al. 2012](#), [Upadhyaya et al. 1992](#)

Number of Variants: 2 <http://bit.ly/2d3R1ec>

Variant Region ID	Type	Number of Variant Calls	Study ID	Organism	Clinical Assertion	Location	Genes in region
nsv997234	copy number 1 variation	1	nstd45	human	Pathogenic	NCBI36 (hg18) chr17: 26,446,071-26,728,821 , GRCh37 (hg19) chr17: 29,421,945-29,704,695 , GRCh38 (hg38) chr17: 31,094,927-31,377,677	EVI2A , EVI2B , 2 more genes
nsv491537	copy number 1 variation	1	nstd45	human	Pathogenic	NCBI36 (hg18) chr17: 26,186,948-27,242,780 , GRCh37 (hg19) chr17: 29,162,822-30,218,667 , GRCh38 (hg38) chr17: 30,835,804-15 more	MIR365B , TRT , CGT4-1 , 15 more

Links to Other Resources

ClinGen: NF1
 GENE: 4763
 GeneReviews: NBK1109
 OMIM: 162200
 OMIM: 613113
 PubMed: 1302608
 PubMed: 1757093
 Overlapping Genes

Source: NCBI

Tabs underneath the summary provide other relevant details. Specifically, the “Genome View” (D) presents the variant graphically by in context with genes mapped to the region. The “Variant Region Details and Evidence” tab (E) provides detailed genomic coordinates on different genome builds. The “Clinical Assertions” tab (F) provides information on the phenotype and clinical interpretation when data is available.

Genome View (D) Variant Region Details and Evidence Validation Information Clinical Assertions Genotype Information

Genome View

Select assembly: [GRCh38 \(hg38\): Chr17](#)

Overlapping variant regions from other studies: 548 SVs from 42 studies. See in: [genome view](#)

Remapped (Score: Perfect): [31,094,927 - 31,377,677](#)

NC_000017.11: 31M..31M (339Kbp) Find:

SV-Anon3 - Segment Map

dbVar ClinGen Curated Dosage Sensitivity Map (nstd45) Pathogenic

nsv997234 (+1)

dbVar ClinGen Curated Dosage Sensitivity Map (nstd45) Benign

Genes

NF1
 LOC105371722
 XR_9346571
 MIR4733

LOC101927057
 XR_243788.4
 EVI2B
 EVI2A
 AKAP1
 NP_002535.3: oligod...
 NM_002544.4
 RAB11FIP

Variant Region Placement Information

Variant Region ID	Placement Type	Score	Assembly	Assembly Unit	Reciprocity	Sequence ID	Chr	Inner Start	Inner Stop
nsv997234	Remapped	Perfect	GRCh38 (hg38)	Primary Assembly	First Pass	NC_000017.11	Chr17	31,094,927	31,377,677
nsv997234	Submitted genomic		GRCh37 (hg19)			NC_000017.10	Chr17	29,421,945	29,704,695
nsv997234	Remapped	Perfect	NCBI36 (hg18)	Primary Assembly	First Pass	NC_000017.9	Chr17	26,446,071	26,728,821

Variant Call Information

Variant Call ID	Type	Method	Analysis	Subject Phenotype	Clinical	Copy number
nsv3442573	copy number loss	Curated	Curated	Neurofibromatosis, type 1	Pathogenic	1

Variant Call Placement Information

Variant Call ID	Placement Type	Score	HGVS	Assembly	Assembly Unit	Reciprocity	Sequence ID	Chr	Inner Start	Inner Stop
nsv3442573	Remapped	Perfect	NC_000017.11:g.(?_31094927)(31377677_?)del	GRCh38 (hg38)	Primary Assembly	First Pass	NC_000017.11	Chr17	31,094,927	31,377,677
nsv3442573	Submitted genomic		NC_000017.10:g.(?_29421945)(29704695_?)del	GRCh37 (hg19)			NC_000017.10	Chr17	29,421,945	29,704,695
nsv3442573	Remapped		NC_000017.9:g.(?_26446071)(26728821_?)del	NCBI36 (hg18)			NC_000017.9	Chr17	26,446,071	26,728,821

Clinical Assertions

Variant Call ID	HGVS	Type	Subject Phenotype	Clinical Interpretation	Copy number
nsv3442573	NC_000017.10:g.(?_29421945)(29704695_?)del	copy number loss	Neurofibromatosis, type 1	Pathogenic	1

References

- Lappalainen I, et. al. (2013) dbVar and DGVA: public archives for genomic structural variation. *Nucleic Acids Res.* 41 (D1):D936-41. [PMID: 23193291](#).
- Sneddon TP and Church DM. (2012) Online resources for genomic structural variation. *Methods Mol Biol.* 838:273-89. [PMID: 22228017](#).
- dbVar help documentation: www.ncbi.nlm.nih.gov/dbvar/content/help/